

(19)

JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 08241325 A

(43) Date of publication of application: 17.09.96

(51) Int. Cl.

G06F 17/30

H03M 7/36

(21) Application number: 07044565

(71) Applicant: MATSUSHITA ELECTRIC IND CO LTD

(22) Date of filing: 03.03.95

(72) Inventor: SHIOMI RYUICHI

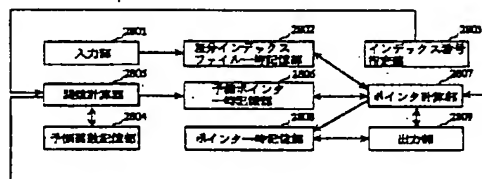
(54) ELECTRONIC DICTIONARY AND ITS PRODUCTION, AND INDEX COMPRESSING AND EXPANDING DEVICE

(57) Abstract:

PURPOSE: To find the storage position of a data element from an inputted difference index file and retrieve the data element.

CONSTITUTION: An input part 2801 receives the difference index file containing difference indexes and stores it in a difference index file temporary storage part 2802. A function calculation part 2805 is informed of an index number from an index number specification part 2803, obtains a specific prediction function from a prediction function storage part 2804, and calculates and stores a prediction pointer corresponding to the index number in a prediction pointer temporary storage part 2806. A pointer calculation part 2807 calculates the sum of the value of a difference index corresponding to the index number stored in the difference index file temporary storage part and the value of the prediction pointer stored in the prediction pointer temporary storage part and stores it in a pointer temporary storage part 2808. An output part 2809 reads the value of the pointer out of the pointer temporary storage part and outputs it.

COPYRIGHT: (C)1996,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-241325

(43) 公開日 平成8年(1996)9月17日

(51) Int.Cl.⁸

G 0 6 F 17/30

H 0 3 M 7/36

識別記号

庁内整理番号

9194-5L

9382-5K

F I.

G 0 6 F 15/40

H 0 3 M 7/36

技術表示箇所

3 7 0 J

審査請求 未請求 請求項の数10 O L (全 21 頁)

(21) 出願番号 特願平7-44565

(22) 出願日 平成7年(1995)3月3日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 塩見 隆一

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

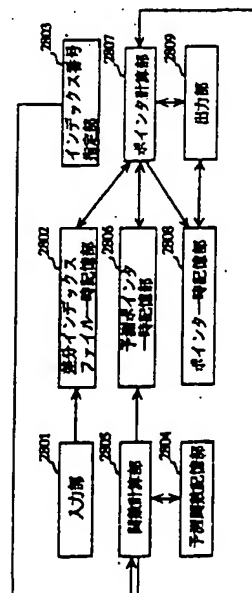
(74) 代理人 弁理士 中島 司朗

(54) 【発明の名称】 電子辞書及びその製造方法並びにインデックス圧縮・伸長装置

(57) 【要約】

【目的】 入力された差分インデックスファイルからデータ要素の格納位置を求めてデータ要素の検索を可能とする。

【構成】 入力部2801は、差分インデックスを格納した差分インデックスファイルの入力を受け、差分インデックスファイル一時記憶部2802に記憶させる。関数計算部2805は、インデックス番号指定部2803からインデックス番号の通知を受け、予測関数記憶部2804から所定の予測関数を取得し、当該インデックス番号に対応する予測ポインタを計算し、予測ポインタ一時記憶部2806に記憶させる。ポインタ計算部2807は、差分インデックスファイル一時記憶部に記憶された当該インデックス番号に対応する差分インデックスの値と予測ポインタ一時記憶部に記憶された予測ポインタの値との和を計算し、ポインタ一時記憶部2808に記憶させる。出力部2809は、ポインタ一時記憶部に記憶されたポインタの値を読み出し出力する。



【特許請求の範囲】

【請求項1】 複数の辞書データ要素からなる辞書データと各辞書データ要素の格納位置を示すインデックスファイルとを辞書媒体内に備える電子辞書であって、インデックスファイルは、各辞書データ要素の格納位置を示すポインタと所定の関数値との差分で表された差分インデックスの集合からなり、各差分インデックスは、辞書データ要素の格納位置を示すポインタの最大値の桁数よりも少ない桁数で表わされていることを特徴とする電子辞書。

【請求項2】 前記電子辞書が複数種類の辞書データを備える電子辞書であって、前記インデックスファイルは更に、所定の関数値を得るための関数情報を含み、前記差分インデックスには、辞書データの種類の識別する識別子を含むことを特徴とする請求項1記載の電子辞書。

【請求項3】 複数の辞書データ要素からなる辞書データと各辞書データ要素の格納位置を示すインデックスファイルとを辞書媒体内に備える電子辞書の製造方法であって、辞書データ格納ステップと、辞書媒体内の辞書データ要素の格納位置を示す各ポインタからなるインデックスファイルを入力する入力ステップと、上記インデックスファイル上における順番を示すインデックス番号を、ポインタの値を予測するため予め所有している関数に代入して予測ポインタを計算する関数計算ステップと、上記インデックスファイル中のポインタと前記関数計算ステップにおいて計算した予測ポインタの差分を計算する差分計算ステップと、前記差分計算ステップにおいて計算した差分を出力する出力ステップと、前記出力ステップにおいて出力された差分を記録するデータ記録ステップとを有することを特徴とする電子辞書の製造方法。

【請求項4】 前記入力ステップと前記関数計算ステップとの間に、上記インデックスファイル中のポインタの値に応じて複数の予測関数から一つの関数を選択する関数選択ステップと、前記関数計算ステップは、前記関数選択ステップで選択した関数を用いて予測ポインタを計算するサブステップを有し、前記出力ステップは、前記関数選択ステップで選択した関数を識別する関数識別子と上記差分とを組として出力するサブステップが加えられたことを特徴とする請求項3記載の電子辞書の製造方法。

【請求項5】 データを構成する複数データ要素から所

望のデータ要素を検索するためのインデックスを、データ要素のデータ内における位置を示すポインタの桁数よりも少ない桁数に圧縮するためのインデックス圧縮装置であって、

データ中におけるデータ要素の位置を示す各ポインタからなるインデックスファイルを入力する入力手段と、入力されたインデックスファイルを一時記憶するインデックスファイル一時記憶手段と、

各ポインタの値を予測する関数を記憶する予測関数記憶手段と、

前記予測関数記憶手段から関数を取得し、インデックスファイルの順番を示すインデックス番号を該関数に代入して予測ポインタを計算する関数計算手段と、

前記関数計算手段が計算した予測ポインタを一時記憶する予測ポインタ一時記憶手段と、

前記インデックスファイル一時記憶手段に記憶されている各ポインタと前記予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算する差分計算手段と、

前記差分計算手段が計算した差分を出力する出力手段とを備えることを特徴とするインデックス圧縮装置。

【請求項6】 前記予測関数記憶手段は、複数の関数を記憶し、

前記インデックスファイル一時記憶手段に記憶されている各ポインタの値に応じて前記予測関数記憶手段が記憶している複数の関数から一の関数を選択する関数選択手段を含み、

前記関数計算手段は、前記関数選択手段が選択した関数を前記予測関数記憶手段から取得し、

前記出力手段は、差分と前記関数選択手段が選択した関数を識別する関数識別子とを組として出力することを特徴とする請求項5記載のインデックス圧縮装置。

【請求項7】 前記予測関数記憶手段に替えて、各ポインタの値を予測する1次関数の傾きと切片との値を記憶する定数記憶手段を設け、

前記関数計算手段は、前記定数記憶手段に記憶されている傾きと切片との値を取得して、この傾きの値にインデックス番号を乗算し、その値に切片の値を加えた予測ポインタを計算することを特徴とする請求項5記載のインデックス圧縮装置。

【請求項8】 データを構成する複数データ要素から所望のデータ要素を検索するためのインデックスを、データ要素のデータ内における位置を示すポインタの桁数よりも少ない桁数に圧縮するためのインデックス圧縮装置であって、

データ中におけるデータ要素の位置を示す各ポインタからなるインデックスファイルを入力する入力手段と、

入力されたインデックスファイルを一時記憶するインデックスファイル一時記憶手段と、

前記インデックスファイル一時記憶手段に記憶されてい

る最後のポインタとインデックスファイルの先頭ポインタからの順番を示す最後のインデックス番号とから各ポインタの値を予測する関数を作成する第1予測関数作成手段と、

前記第1予測関数作成手段が作成した関数を記憶する第1予測関数記憶手段と、

前記第1予測関数記憶手段から関数を取得し、インデックス番号を該関数に代入して予測ポインタを計算する第1関数計算手段と、

前記第1関数計算手段が計算した予測ポインタを一時記憶する予測ポインタ一時記憶手段と、

前記インデックスファイル一時記憶手段に記憶されている各ポインタと前記予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算する第1差分計算手段と、

前記差分計算手段が計算した差分を一時記憶する差分一時記憶手段と、

差分の下限と上限とを示す差分範囲を記憶している差分範囲記憶手段と、

前記差分一時記憶手段に記憶されている各差分が前記差分範囲記憶手段に記憶されている差分範囲内であるか否かを判定する差分判定手段と、

前記差分判定手段が差分範囲内と判定したときに前記差分一時記憶手段に記憶されている差分と前記第1予測関数記憶手段に記憶されている関数とを出力する第1出力手段と、

前記差分判定手段が差分範囲内でないと判定したときに、前記インデックスファイル一時記憶手段に記憶されているポインタの増分の変化値を全ポインタについて計算し、その変化値が所定の値を超えるポインタを境界としてインデックスファイルの区間を分割する区間分割手段と、

前記区間分割手段で分割された各区間で各ポインタの値を予測する異なる関数を作成する第2予測関数作成手段と、

前記区間分割手段で分割された区間と前記第2予測関数作成手段で作成された関数とを組にした一覧表を作成して一時記憶する一覧表作成一時記憶手段と、

一覧表作成一時記憶手段に記憶されている関数を取得し、インデックス番号を該関数に代入して予測ポインタを計算する第2関数計算手段と、

前記第2関数計算手段が計算した予測ポインタを一時記憶する第2予測ポインタ一時記憶手段と、

前記インデックスファイル一時記憶手段に記憶されている各ポインタと前記第2予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算する第2差分計算手段と、

前記第2差分計算手段が計算した差分を一時記憶する第2差分一時記憶手段と、

前記一覧表作成一時記憶手段に記憶されている一覧表と

第2差分一時記憶手段に記憶されている差分とを出力する第2出力手段とを備えることを特徴とするインデックス圧縮装置。

【請求項9】 複数のデータ要素からなるデータファイル上の各データ要素のポインタの値を予測する関数を記憶する予測関数記憶手段と、

各データ要素のポインタと前記予測関数記憶手段に記憶されている関数で予測される対応する予測ポインタとの差分からなる差分インデックスファイルを入力する入力手段と、

前記入力手段で入力された差分インデックスファイルを一時記憶する差分インデックスファイル一時記憶手段と、

前記差分一時記憶手段に記憶されている差分の順番であるインデックス番号を指定するインデックス番号指定手段と、

前記インデックス番号指定手段で指定されたインデックス番号を前記予測関数記憶手段に記憶されている関数に代入して予測ポインタの値を計算する関数計算手段と、

前記関数計算手段で計算された予測ポインタの値と前記差分インデックスファイル一時記憶手段に記憶されている差分とを加算してポインタの値を計算するポインタ計算手段と、

前記ポインタ計算手段で計算したポインタの値を出力する出力手段とを備えることを特徴とするインデックス伸長装置。

【請求項10】 前記予測関数記憶手段は、

複数のデータ要素からなるデータファイル上の各データ要素のポインタを予測する複数の関数を入力する関数入力部と、

前記予測関数入力部から入力された関数を記憶する予測関数記憶部とを備え、

前記関数計算手段は、

前記インデックス番号指定手段で指定されたインデックス番号に基づいて前記予測関数記憶部に記憶されている複数の関数から一の関数を選択する選択部と、

前記選択部で選択された関数にインデックス番号を代入して予測ポインタを計算する関数計算部とを備えることを特徴とする請求項9記載のインデックス伸長装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、電子辞書に関し、特に電子辞書のデータ要素の格納位置を示すインデックスの圧縮方法及び装置並びに伸長装置に関する。

【0002】

【従来の技術】 従来の電子英和辞書の一例を図35に示す。この電子英和辞書は、英語見出し語とその訳語とからなる辞書データ要素を複数格納している英語辞書データ3501と、英語辞書データ3501の各辞書データ要素の格納位置を示すポインタ3502の集合したイン

デックスファイル3503とからなる。

【0003】英語辞書データに格納されている辞書データ要素のサイズ(大きさ)は一定ではなく、辞書データ要素は隙間無く格納されている。インデックスファイル3503は、この辞書データ要素を高速に検索するため、各辞書データ要素の先頭格納位置を示すポインタ3502を格納している。このポインタは、同じサイズ(桁数)で表現されている。このサイズをSとすると、T番目の辞書データ要素の先頭格納位置を知るためには、インデックスファイル3503の先頭位置からS*T番目からSだけ読めば、T番目の辞書データ要素の先頭格納位置を得ることができる。これによって、辞書データ要素を高速に検索することができる。

【0004】また、この電子英和辞書を入力された英語キーワードを用いて検索することも可能である。この検索は、一般的に表探索に分類されており、「データ構造とプログラム技法」(萩原宏、西原清一著、オーム社出版、P172~176)に順探索法と2分探索法とが記載されている。この順探索法では、各辞書データ要素の英語見出し語と英語キーワードとを辞書データ要素の格納順に比較するものである。この際、インデックスファイル3502があれば、各辞書データ要素の先頭格納位置を容易に取得することができ、高速に各辞書データ要素の英語見出し語と英語キーワードとを比較することができる。

【0005】

【発明が解決しようとする課題】ところで、上記の従来の電子英和辞書のインデックスファイル3503は、英和辞書データ3501上でのデータ要素の先頭格納位置を一定のサイズで表現するため、インデックスファイル3503中の最大の先頭格納位置を表現できるサイズに統一する。この結果、インデックスファイル3503の記憶容量は大きくなる。

【0006】また、インデックスファイル3503がなければ、各データ要素の先頭格納位置を見つけるため英和辞書データ3501を先頭から順に調べる必要があり、多くの検索時間を要する。一方、インデックスファイル3503を設けずに、各辞書データ要素のサイズを一定とした英和辞書データ3501とすることによって、高速に検索することは可能であるが、インデックスファイル3503の容量が大きくなるという同じ理由で英和辞書データ3501の容量が非常に大きくなってしまふ。

【0007】本発明は、上記課題に鑑み、省記憶容量にしたインデックスファイルを含む電子辞書を提供することを目的とする。本発明の他の目的は、従来のインデックスファイルの容量を圧縮するインデックス圧縮方法の提供である。本発明の他の目的は、インデックス圧縮装置の提供である。

【0008】本発明の更に他の目的は、圧縮されたイン

デックスを復元するインデックス伸長装置の提供である。

【0009】

【課題を解決するための手段】上記課題を解決するため、請求項1に係る発明においては、複数の辞書データ要素からなる辞書データと各辞書データ要素の格納位置を示すインデックスファイルとを辞書媒体内に備える電子辞書であって、インデックスファイルは、各辞書データ要素の格納位置を示すポインタと所定の関数値との差分で表された差分インデックスの集合からなり、各差分インデックスは、辞書データ要素の格納位置を示すポインタの最大値の桁数よりも少ない桁数で表わされていることを特徴としている。

【0010】請求項2に係る発明においては、前記電子辞書が複数種類の辞書データを備える電子辞書であって、前記インデックスファイルは更に、所定の関数値を得るための関数情報を含み、前記差分インデックスには、辞書データの種類の識別する識別子を含むことを特徴としている。請求項3に係る発明においては、複数の辞書データ要素からなる辞書データと各辞書データ要素の格納位置を示すインデックスファイルとを辞書媒体内に備える電子辞書の製造方法であって、辞書データ格納ステップと、辞書媒体内の辞書データ要素の格納位置を示す各ポインタからなるインデックスファイルを入力する入力ステップと、上記インデックスファイル上における順番を示すインデックス番号を、ポインタの値を予測するため予め所有している関数に代入して予測ポインタを計算する関数計算ステップと、上記インデックスファイル中のポインタと前記関数計算ステップにおいて計算した予測ポインタの差分を計算する差分計算ステップと、前記差分計算ステップにおいて計算した差分を出力する出力ステップと、前記出力ステップにおいて出力された差分を記録するデータ記録ステップとを有することを特徴としている。

【0011】請求項4に係る発明においては、前記入力ステップと前記関数計算ステップとの間に、上記インデックスファイル中のポインタの値に応じて複数の予測関数から一つの関数を選択する関数選択ステップと、前記関数計算ステップは、前記関数選択ステップで選択した関数を用いて予測ポインタを計算するサブステップを有し、前記出力ステップは、前記関数選択ステップで選択した関数を識別する関数識別子と上記差分とを組として出力するサブステップが加えられたことを特徴としている。

【0012】請求項5に係る発明においては、データを構成する複数データ要素から所望のデータ要素を検索するためのインデックスを、データ要素のデータ内における位置を示すポインタの桁数よりも少ない桁数に圧縮するためのインデックス圧縮装置であって、データ中におけるデータ要素の位置を示す各ポインタからなるインデ

ックスファイルを入力する入力手段と、入力されたインデックスファイルを一時記憶するインデックスファイル一時記憶手段と、各ポインタの値を予測する関数を記憶する予測関数記憶手段と、前記予測関数記憶手段から関数を取得し、インデックスファイルの順番を示すインデックス番号を該関数に代入して予測ポインタを計算する関数計算手段と、前記関数計算手段が計算した予測ポインタを一時記憶する予測ポインタ一時記憶手段と、前記インデックスファイル一時記憶手段に記憶されている各ポインタと前記予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算する差分計算手段と、前記差分計算手段が計算した差分を出力する出力手段とを備えることを特徴としている。

【0013】請求項6に係る発明においては、前記予測関数記憶手段は、複数の関数を記憶し、前記インデックスファイル一時記憶手段に記憶されている各ポインタの値に応じて前記予測関数記憶手段が記憶している複数の関数から一の関数を選択する関数選択手段を含み、前記関数計算手段は、前記関数選択手段が選択した関数を前記予測関数記憶手段から取得し、前記出力手段は、差分と前記関数選択手段が選択した関数を識別する関数識別子とを組として出力することを特徴としている。

【0014】請求項7に係る発明においては、前記予測関数記憶手段に替えて、各ポインタの値を予測する1次関数の傾きと切片との値を記憶する定数記憶手段を設け、前記関数計算手段は、前記定数記憶手段に記憶されている傾きと切片との値を取得して、この傾きの値にインデックス番号を乗算し、その値に切片の値を加えた予測ポインタを計算することを特徴としている。

【0015】請求項8に係る発明においては、データを構成する複数データ要素から所望のデータ要素を検索するためのインデックスを、データ要素のデータ内における位置を示すポインタの桁数よりも少ない桁数に圧縮するためのインデックス圧縮装置であって、データ中におけるデータ要素の位置を示す各ポインタからなるインデックスファイルを入力する入力手段と、入力されたインデックスファイルを一時記憶するインデックスファイル一時記憶手段と、前記インデックスファイル一時記憶手段に記憶されている最後のポインタとインデックスファイルの先頭ポインタからの順番を示す最後のインデックス番号とから各ポインタの値を予測する関数を作成する第1予測関数作成手段と、前記第1予測関数作成手段が作成した関数を記憶する第1予測関数記憶手段と、前記第1予測関数記憶手段から関数を取得し、インデックス番号を該関数に代入して予測ポインタを計算する第1関数計算手段と、前記第1関数計算手段が計算した予測ポインタを一時記憶する予測ポインタ一時記憶手段と、前記インデックスファイル一時記憶手段に記憶されている各ポインタと前記予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算する第1差

分計算手段と、前記差分計算手段が計算した差分を一時記憶する差分一時記憶手段と、差分の下限と上限とを示す差分範囲を記憶している差分範囲記憶手段と、前記差分一時記憶手段に記憶されている各差分が前記差分範囲記憶手段に記憶されている差分範囲内であるか否かを判定する差分判定手段と、前記差分判定手段が差分範囲内と判定したときに前記差分一時記憶手段に記憶されている差分と前記第1予測関数記憶手段に記憶されている関数とを出力する第1出力手段と、前記差分判定手段が差分範囲内でないと判定したときに、前記インデックスファイル一時記憶手段に記憶されているポインタの増分の変化値を全ポインタについて計算し、その変化値が所定の値を超えるポインタを境界としてインデックスファイルの区間を分割する区間分割手段と、前記区間分割手段で分割された各区間で各ポインタの値を予測する異なる関数を作成する第2予測関数作成手段と、前記区間分割手段で分割された区間と前記第2予測関数作成手段で作成された関数とを組にした一覧表を作成して一時記憶する一覧表作成一時記憶手段と、一覧表作成一時記憶手段に記憶されている関数を取得し、インデックス番号を該関数に代入して予測ポインタを計算する第2関数計算手段と、前記第2関数計算手段が計算した予測ポインタを一時記憶する第2予測ポインタ一時記憶手段と、前記インデックスファイル一時記憶手段に記憶されている各ポインタと前記第2予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算する第2差分計算手段と、前記第2差分計算手段が計算した差分を一時記憶する第2差分一時記憶手段と、前記一覧表作成一時記憶手段に記憶されている一覧表と第2差分一時記憶手段に記憶されている差分とを出力する第2出力手段とを備えることを特徴としている。

【0016】請求項9に係る発明においては、複数のデータ要素からなるデータファイル上の各データ要素のポインタの値を予測する関数を記憶する予測関数記憶手段と、各データ要素のポインタと前記予測関数記憶手段に記憶されている関数で予測される対応する予測ポインタとの差分からなる差分インデックスファイルを入力する入力手段と、前記入力手段で入力された差分インデックスファイルを一時記憶する差分インデックスファイル一時記憶手段と、前記差分一時記憶手段に記憶されている差分の順番であるインデックス番号を指定するインデックス番号指定手段と、前記インデックス番号指定手段で指定されたインデックス番号を前記予測関数記憶手段に記憶されている関数に代入して予測ポインタの値を計算する関数計算手段と、前記関数計算手段で計算された予測ポインタの値と前記差分インデックスファイル一時記憶手段に記憶されている差分とを加算してポインタの値を計算するポインタ計算手段と、前記ポインタ計算手段で計算したポインタの値を出力する出力手段とを備えることを特徴としている。

【0017】請求項10に係る発明においては、前記予測関数記憶手段は、複数のデータ要素からなるデータファイル上の各データ要素のポイントを予測する複数の関数を入力する関数入力部と、前記予測関数入力部から入力された関数を記憶する予測関数記憶部とを備え、前記関数計算手段は、前記インデックス番号指定手段で指定されたインデックス番号に基づいて前記予測関数記憶部に記憶されている複数の関数から一の関数を選択する選択部と、前記選択部で選択された関数にインデックス番号を代入して予測ポイントを計算する関数計算部とを備えることを特徴としている。

【0018】

【作用】上記構成により、請求項1の発明において、インデックスファイルは、各辞書データ要素の格納位置を示すポイントと各辞書データのポイントを近似する所定の関数値との差分で表された差分インデックスの集合からなる。各差分インデックスの2進数表示の桁数は辞書データ要素の格納位置を示すポイントの2進数表示の桁数の最大値よりも少ない桁数で表されるので、電子辞書は、省記憶容量としたインデックスファイルを有する。

【0019】請求項2の発明において、インデックスファイルは所定の関数値を得るための関数情報を含むので、差分インデックスからポイントを容易に復元でき、差分インデックスが辞書データの種類の識別子を含むので、複数の系列の辞書データ（基本辞書データとユーザ辞書データ）を含む電子辞書にも、差分インデックスファイルを適用できる。

【0020】請求項3の発明において、辞書データ格納ステップは、複数の辞書データ要素を格納する。入力ステップは、辞書媒体内の辞書データ要素の格納位置を示す各ポイントからなるインデックスファイルを入力する。関数計算ステップは、上記インデックスファイル上における先頭ポイントからの順番を示すインデックス番号を、ポイントの値を予測するため予め所有している関数に代入して予測ポイントを計算する。差分計算ステップは、上記インデックスファイル中のポイントと前記関数計算ステップにおいて計算した予測ポイントの差分を計算する。これによって、差分の2進数表示の桁数は、元のポイントの2進数表示の桁数の最大値よりも少ない桁数で表現される。出力ステップは、前記差分計算ステップにおいて計算した差分を出力する。データ記録ステップは、前記出力ステップにおいて出力された差分を小さな記憶容量で記録する。

【0021】請求項4の発明において、関数選択ステップは、上記インデックスファイル中のポイントの値に応じて複数の予測関数（複数の系列の辞書データ要素に対応する）から一つの関数を選択する。前記関数計算ステップは、前記関数選択ステップで選択した関数を用いて予測ポイントを計算する。前記出力ステップは、前記関数選択ステップで選択した関数を識別する関数識別子と

上記差分とを組として出力する。これによって、複数の系列の辞書データの種類にも対応することができる。

【0022】請求項5の発明において、入力手段は、データ中におけるデータ要素の位置を示す各ポイントからなるインデックスファイルを入力すると、インデックスファイル一時記憶手段に一時記憶させる。関数計算手段は、予測関数記憶手段から各ポイントの値を予測する関数を取得し、インデックスファイルの先頭ポイントからの順番を示すインデックス番号を該関数に代入して予測ポイントを計算し、予測ポインター一時記憶手段に、一時記憶させる。差分計算手段は、インデックスファイル一時記憶手段に記憶されている各ポイントと前記予測ポインター一時記憶手段に記憶されている対応する予測ポイントとの差分を計算する。この差分の2進数表示の桁数は、元のポイントの2進数表示の最大桁数に比べて小さい。出力手段は、前記差分計算手段が計算した差分を出力する。

【0023】請求項6の発明において、関数選択手段は、インデックスファイル一時記憶手段に記憶されている各ポイントの2進数表示の所定桁の値に応じて前記予測関数記憶手段が記憶している複数の関数から一の関数を選択する。関数計算手段は、関数選択手段が選択した関数を予測関数記憶手段から取得する。これによって、予測ポイントの計算に用いる関数を使い分けることができる。出力手段は、差分と前記関数選択手段が選択した関数を識別する関数識別子とを組として出力する。

【0024】請求項7の発明において、関数計算手段は、定数記憶手段に記憶されている各ポイントの値を予測する1次関数の傾きと切片との値を取得して、この傾きの値にインデックス番号を乗算し、その値に切片の値を加えた予測ポイントを計算する。請求項8の発明において、入力手段は、データ中におけるデータ要素の位置を示す各ポイントからなるインデックスファイルを入力すると、インデックスファイル一時記憶手段に記憶させる。第1予測関数作成手段は、インデックスファイル一時記憶手段に記憶されている最後のポイントの値を分子とし、インデックスファイルの先頭ポイントからの順番を示す最後のインデックス番号の値を分母として各ポイントの値を予測する一次関数を作成し、第1予測関数記憶手段に記憶させる。第1関数計算手段は、第1予測関数記憶手段から関数を取得し、インデックス番号を該関数に代入して予測ポイントを計算し、予測ポインター一時記憶手段に一時記憶させる。第1差分計算手段は、インデックスファイル一時記憶手段に記憶されている各ポイントと予測ポインター一時記憶手段に記憶されている対応する予測ポイントとの差分を計算し、差分一時記憶手段に一時記憶させる。差分判定手段は、差分一時記憶手段に記憶されている各差分が差分範囲記憶手段に記憶されている差分の下限と上限とを示す差分範囲内であるか否かを判定する。第1出力手段は、差分判定手段が差分範

図内と判定したときに即ち、差分を所定ビット内で記述できるときに差分一時記憶手段に記憶されている差分と第1予測関数記憶手段に記憶されている関数とを出力する。区間分割手段は、差分判定手段が差分範囲内でないと判定したときに、インデックスファイル一時記憶手段に記憶されているポインタの増分の変化値(2階差分の値)を全ポインタについて計算し、その変化値が所定の値を超えるポインタを境界としてインデックスファイルの区間を分割する。第2予測関数作成手段は、区間分割手段で分割された各区間で各ポインタの値を予測する異なる関数を作成する。一覧表作成一時記憶手段は、区間分割手段で分割された区間と第2予測関数作成手段で作成された関数とを組にした一覧表を作成して一時記憶する。第2関数計算手段は、一覧表作成一時記憶手段に記憶されている関数を取得し、インデックス番号を該関数に代入して予測ポインタを計算し、第2予測ポインタ一時記憶手段に一時記憶させる。第2差分計算手段は、インデックスファイル一時記憶手段に記憶されている各ポインタと第2予測ポインタ一時記憶手段に記憶されている対応する予測ポインタとの差分を計算し、第2差分一時記憶手段に記憶させる。予測ポインタの計算精度が向上しているため差分の値は小さくなっている。第2出力手段は、一覧表作成一時記憶手段に記憶されている一覧表と第2差分一時記憶手段に記憶されている差分とを出力する。

【0025】請求項9の発明において、入力手段は、各データ要素のポインタと予測関数記憶手段に記憶されている複数のデータ要素からなるデータファイル上の各データ要素のポインタの値を予測する関数で予測される対応する予測ポインタとの差分からなる差分インデックスファイルを出力する。差分インデックスファイルを差分インデックスファイル一時記憶手段に一時記憶させる。インデックス番号指定手段は、差分一時記憶手段に記憶されている差分の先頭からの順番であるインデックス番号をデータ要素検索のため指定する。関数計算手段は、インデックス番号指定手段で指定されたインデックス番号を予測関数記憶手段に記憶されている関数に代入して予測ポインタの値を計算する。ポインタ計算手段は、関数計算手段で計算された予測ポインタの値と差分インデックスファイル一時記憶手段に記憶されている差分とを加算してポインタの値を計算する。出力手段は、ポインタ計算手段で計算して復元されたポインタの値を出力する。これによって、省容量の差分インデックスファイルからポインタを求めることができる。

【0026】請求項10の発明において、予測関数記憶手段の関数入力部は、複数のデータ要素からなるデータファイル上の各データ要素のポインタを予測する複数の関数を入力すると、予測関数記憶部に記憶させる。関数計算手段の選択部は、インデックス番号指定手段で指定されたインデックス番号に基づいて前記予測関数記憶部

に記憶されている複数の関数から一の関数を選択する。関数計算部は、選択部で選択された関数にインデックス番号を代入して予測ポインタを計算する。

【0027】

【実施例】以下、本発明に係るインデックス圧縮装置を実施例に基づいて、説明する。

(第1実施例) 図1は、本発明に係るインデックス圧縮装置の第1実施例の構成図である。このインデックス圧縮装置は、入力部101と、インデックスファイル一時記憶部102と、予測関数記憶部103と、関数選択部104と、関数識別子一時記憶部105と、関数計算部106と、予測ポインタ一時記憶部107と、差分計算部108と、差分一時記憶部109と、出力部110とを備える。

【0028】入力部101は、ポインタを格納したインデックスファイルの入力を受けると、インデックスファイル一時記憶部102に記憶させるとともに、関数選択部104に起動指示をする。インデックスファイル一時記憶部102は、高速半導体メモリ等などで実現され、入力部101から入力されたインデックスファイルを一時記憶する。

【0029】図2は、インデックスファイル一時記憶部102に一時記憶されているインデックスファイル210の内容を示す。インデックスファイル210は、例えば英和辞書の基本辞書データ220とユーザ辞書データ230との各データの格納位置211と基本辞書データ220とユーザ辞書データ230とのいずれのデータであるかを示す辞書フラグ212とからなるポインタを有している。なお、英和辞書に基本辞書データ220とユーザ辞書データ230との2種が含まれるとしたので辞書フラグ212を含めたけれども、一種の辞書である場合には、辞書フラグ212は省略できる。

【0030】インデックスファイル210の左方に記録されているインデックス番号213は、入力部101がインデックスファイル一時記憶部102にインデックスファイルを記憶させる際に各ポインタの記憶順に0、1、2、…と番号を付したものであり、最後のポインタのインデックス番号が「1523」であるので、基本辞書データ220とユーザ辞書データ230とのデータ要素の総数が「1524」であることを示している。

【0031】ここで基本辞書データ220、ユーザ辞書データ230の内容について簡単に説明する。基本辞書データ220等は、磁気ディスク等で実現されるデータファイル格納部(図示せず)に格納されている。基本辞書データ220、ユーザ辞書データ230中の□221、222、231、232は、英語見出し語と訳語との終端を表す終端記号である。基本辞書データ220の最初のデータ要素「a□1つの、1人の、ある□」は、英語文字を1バイト、日本語文字を2バイト、終端記号を1バイトで表現するので、23バイトである。次のデ

ータ要素「a1ロー流の、極上の口」は18バイトである。これらのデータ要素は、データファイル記憶部に隙間無く格納されている。

【0032】辞書フラグ212は、その値が「0」のときそのデータが基本辞書データ220に含まれ、その値が「1」のときそのデータがユーザ辞書データ230に含まれることを示している。格納位置211は、各データ要素の基本辞書データ220又はユーザ辞書データ230の先頭位置からのバイト数で表現される。したがって、インデックス番号0の最初のポインタは、辞書フラグ212の値が「0」、格納位置211の値が「0」で基本辞書データ220の「0」バイト目に格納されている最初のデータ要素「a1つの、1人の、ある口」を示している。またインデックス番号2のポインタは、辞書フラグ212の値が「1」、格納位置211の値が「0」でユーザ辞書データ230の「0」バイト目に記憶されている最初のデータ要素「AAA口スリーA口」を示している。インデックス番号3のポインタは、辞書フラグ212の値が「0」で格納位置211の値が「41」で基本辞書データ220の「41」バイト目に格納されているデータ要素「a back口あとへ、うしろへ口」を示している。これらの関係は図2の矢印241～246等で示している。

【0033】なお、インデックスファイル210は、高速半導体メモリ上では、図3に示すバイナリ表現されたインデックスファイル301の形態で記憶されている。インデックスファイル302は、これと等価な16進表現されたものである。インデックスファイル301の上位バイトの最上位ビットは辞書フラグ312を示し、上位バイトの下位7ビットと下位バイトの合計15ビットで格納位置311を示している。即ち、本実施例で入力されるインデックスファイル210の各ポインタは2バイトで表現されている。

【0034】図4は、入力部101から図2に示したインデックスファイル210が入力された直後のインデックスファイル一時記憶部102と予測ポインタ一時記憶部107と差分一時記憶部109と関数識別子一時記憶部105との記憶内容を示している。インデックスファイル一時記憶部102以外は、まだ何も記憶していない。

【0035】予測関数記憶部103は、磁気ディスク、高速半導体メモリ等で実現され、インデックスファイル一時記憶部102に記憶されたポインタの値を予測する予測ポインタを計算するための予測関数を1つ以上記憶する。予測関数記憶部103は、図5に示す第1予測関数501と第2予測関数502とを記憶している。Xを入力変数、Yを出力変数とすると第1予測関数501は $Y=21X$ であり、第2予測関数502は $Y=0$ である。

【0036】関数選択部104は、予測関数記憶部10

3に記憶された予測関数を選択する。関数選択部104は、入力部101から起動指示を受けると、インデックスファイル一時記憶部102に記憶されている各ポインタの辞書フラグ212を順に読み出し、関数識別子一時記憶部105に辞書フラグの値が「0」のとき第1予測関数501を識別する識別子「0」を記憶させ、辞書フラグの値が「1」のとき第2予測関数502を識別する識別子「1」を記憶させる。インデックス番号213の「1523」までが終了すると、関数選択部104は、関数計算部106に起動指示をする。

【0037】関数識別子一時記憶部105は、高速半導体メモリ等で実現され、関数選択部104で選択された関数に対応する識別子を一時記憶する。関数選択部104によって識別子が記憶させられると、関数識別子一時記憶部105は、図6に示すような記憶内容となる。関数計算部106は、関数識別子一時記憶部105に記憶された識別子を読み出し、使用する関数を予測関数記憶部103から取得し、インデックスファイル一時記憶部102に記憶されているポインタに対応するインデックス番号を取得した関数に代入して、ポインタの値を計算して予測ポインタを得る。関数計算部106は関数選択部104から起動指示を受けると、インデックス番号213の「0」から順に「1523」まで予測ポインタの計算をする。例えば、インデックス番号213の「0」の予測ポインタは、関数識別子一時記憶部105に記憶されている識別子「0」を読み出し、予測関数記憶部103に記憶されている識別子「0」で識別される第1予測関数501の $Y=21X$ を取得し、入力変数Xにインデックス番号213の「0」を代入して計算し、 $Y=21 \times 0 = 0$ から予測ポインタの値「0」を計算し、予測ポインタ一時記憶部107に記憶させる。予測ポインタの値をインデックス番号213の「1523」まで終了すると、差分計算部108に起動指示をする。

【0038】予測ポインタ一時記憶部107は、高速半導体メモリ等で実現され、関数計算部106で計算された予測ポインタの値を一時記憶する。関数計算部106によって予測ポインタの値が全て計算されると、予測ポインタ一時記憶部107は、図7に示すような状態となる。差分計算部108は、インデックス一時記憶部102に記憶されているポインタの値と予測ポインタ一時記憶部107に記憶されている予測ポインタの値の差分を計算する。差分計算部108は、関数計算部106から起動指示を受けると、インデックス番号213の「0」から「1523」まで、順次インデックスファイル一時記憶部102の格納位置211の値と予測ポインタ一時記憶部107の予測ポインタの値とをそれぞれ読み出し、格納位置211の値から予測ポインタの値を減算し、得られた差分を差分一時記憶部109に記憶させる。全ての差分の計算が終了すると、出力部110に起動指示をする。

【0039】差分一時記憶部109は、高速半導体メモリ等で実現され、差分計算部108で計算された差分を一時記憶する。差分計算部108で差分が全て計算されると、差分一時記憶部109は、図8に示すような状態となる。出力部110は、差分計算部108から起動指示を受けると、差分一時記憶部109と関数識別子一時記憶部105とに記憶された差分と識別子とを組にした差分インデックスを記録する記録部（図示せず）に出力する。出力部110の出力が終了すると、差分インデックスは、図9に示すように、1ビットからなる辞書フラグ901と7ビットからなる差分902とを組とした差分インデックスファイル903となる。図10は、本実施例のインデックスファイル圧縮装置によってインデックスファイルの各ポインタの容量を1バイトに圧縮した差分ファイル1001を含む電子辞書の内容を示す。この電子辞書は、英和辞書であり、基本辞書データ1002とユーザ辞書データ1003とこれらのデータ要素の格納位置を示す差分インデックスファイル1001とを有する。

【0040】差分インデックスファイル1001は、ヘッダ情報1004と、各データ要素に対応する辞書フラグ1005と差分1006とからなるポインタとを有する。ヘッダ情報1004は、データ要素のポインタを予測する予測関数の個数である関数個数1007と、予測関数が一次関数である場合の傾き1008と、切片1009と、データ要素の数である差分個数1010と、辞書フラグ1005の記録ビット数であるフラグビット数1011と、差分1006の記録ビット数である差分ビット数1012とを有している。これらの各情報は、2バイトで記録されている。

【0041】各ポインタは、1ビットで記録された辞書フラグと7ビットで記録された差分からなり、各ポインタの左方に記録された数字はインデックス番号1013を示している。なお、この電子辞書を専用の装置で使用する場合には、ヘッダ情報1004は、専用の装置内に組み込まれたプログラムに内蔵しておくことが可能となり、電子辞書の差分インデックスファイルにヘッダ情報1004として記録する必要はない。さらに、電子辞書内に2つの種類の基本辞書データ1002とユーザ辞書データ1003とを含んでいるので辞書フラグ1005を各ポインタに記録したけれども、一種類のデータであるときには、辞書フラグ1005を記録する必要はない。

【0042】次に、本実施例の動作を図11に示すフローチャートを用いて説明する。入力部101からインデックスファイルを入力すると（S1102）、インデックスファイル一時記憶部102に一時記憶される（S1104）。関数選択部104は、インデックスファイル一時記憶部102が記憶しているインデックスファイル中の各ポインタの第1桁の辞書フラグの値が「0」か

「1」か判定し（S1106）、「0」のときは第1予測関数を選択し（S1108）、「1」のときは第2予測関数を選択し（S1110）、対応する識別子を関数識別子一時記憶部105に記憶させる（S1112）。全てのポインタに対応して識別子を関数識別子一時記憶部105に一時記憶させたか否か判定し（S1114）、終了しているときはS1116に移り、まだのときはS1106に戻る。S1116において、関数計算部106は、関数識別子一時記憶部105が記憶する関数の識別子を読み出し、予測関数記憶部103から予測関数を取得する。次に、インデックスファイル一時記憶部102が記憶しているポインタに対応するインデックス番号を取得した関数に代入して予測ポインタを計算し（S1118）、予測ポインタ一時記憶部107に一時記憶させる（S1120）。予測ポインタの計算が全て終了したか否か判定し（S1122）、終了しているときはS1124に移り、まだのときはS1116に戻る。S1124において、差分計算部108は、インデックスファイル一時記憶部102が記憶しているポインタと予測ポインタ一時記憶部107が記憶している予測ポインタとの差分を計算し、これを差分一時記憶部109に一時記憶させる（S1126）。

【0043】最後に、出力部110は、差分一時記憶部109の差分とこれに対応する関数識別子一時記憶部105が記憶している関数の識別子とを組として差分インデックスファイル903を出力して（S1130）、処理を終了する。なお、本実施例の予測関数記憶部103は、一次関数と定数関数を記憶しているが、圧縮対象となるインデックスファイル中のポインタに応じて2次関数や指数関数、切り上げ、切り下げなどの演算を含む関数などでも実施可能である。

【0044】また、差分計算部108は、ポインタの値から予測ポインタの値を減じていたが、予測ポインタの値からポインタの値を減じて実施可能である。なお、本実施例では、英和辞書に2種のデータが含まれている場合についてのインデックスファイルの圧縮について述べたけれども、3種以上のデータが含まれる場合にも、又英和辞書に替えて国語辞典、アドレス帳、データベースのデータ等のインデックスファイルの圧縮についても実施可能なことはいうまでもない。

（第2実施例）図12は、本発明に係るインデックス圧縮装置の第2実施例の構成図である。

【0045】このインデックスファイル圧縮装置は、入力部101と、インデックスファイル一時記憶部102と、差分範囲記憶部1201と、インデックス分割部1202と、予測関数記憶部103と、関数作成部1203と、一覧表作成部1204と、一覧表一時記憶部1205と、関数計算部1206と、予測ポインタ一時記憶部107と、差分計算部108と、差分一時記憶部109と、分割指示部1207と、一覧表出力部1208と

を備える。上記第1実施例の各部と同一の構成を有する各部は、同一の符号を付してその説明を省略する。

【0046】入力部101は、図13に示すようなインデックスファイル1302を入力されると、インデックスファイル一時記憶部102に一時記憶させるとともにインデックス分割部1202に起動指示をする。図13は、インデックスファイル1301とその各ポインタ1302によって格納位置が示される固有名詞辞書データ1303とを示している。

【0047】インデックスファイル1301は、固有名詞辞書データ1303のデータ要素の格納位置を示す各ポインタ1302を格納する。各ポインタ1302には図2のインデックスファイル210同様、インデックス番号1303が振られている。固有名詞辞書データ1304は、各ポインタ1302で格納位置を示されたデータ要素を含んでいる。

【0048】ここでインデックス番号2のポインタは40で、固有名詞辞書データ1304の40バイト目から格納されているデータ要素「岩手県名、地名」を指し示している。図14は、インデックスファイル1301をバイナリ表現したものであり、磁気ディスク上などにはこの形態で格納される。各ポインタは2バイトで表現されており、インデックスファイル1401のポインタは16進数で表現されている。

【0049】このように、図14のインデックスファイル1401はポインタを2バイトで表現しているが、本実施例によりポインタを1バイトで表現する。図15は、入力直後の状態での各一時記憶部102、107、109、1205の記憶内容を表したものである。差分範囲記憶部1201は、磁気ディスクや高速半導体メモリで実現され、差分一時記憶部109に記憶されている差分の上限と下限とを図16に示すように記憶している。ここで、差分の上限を「+127」、差分の下限を「-128」としたのは、「+」、「-」の記号を含めて8ビットの範囲内で差分を記憶しておくためである。

【0050】インデックス分割部1202は、インデックスファイル一時記憶部102で記憶されているインデックスファイルを1つ以上の区間に分割する。入力部101から起動指示を受けると、全区間を1つの区間とし、分割は行わず、インデックスファイル一時記憶部102から最終インデックス番号を取り出し、関数作成部1203に通知する。また、分割指示部1207から分割指示を受けると、インデックスファイル一時記憶部102のインデックス番号1303の「0」のポインタ1302から順次その値を読み出し、連続するポインタ1302間の値の増加値を計算し、更にその隣接する増加値の変化値を計算する（即ち、ポインタ1302の値の2階差分値を計算する）。最後のインデックス番号のポインタ1302までの変化値の計算を終了すると、差分範囲記憶部1201に記憶されている範囲を超えた変化

値に対応するインデックス番号を関数作成部1203に通知する。

【0051】関数作成部1203は、インデックス分割部1202から通知されたインデックス番号1303を境界として、そのインデックス番号1303を区間の最後とし、そのインデックス番号1303に後続するインデックス番号1303があるときは、そのインデックス番号1303の次のインデックス番号1303を区間の最初として、各分割された区間のポインタ1302の値を予測するための関数を作成し、予測関数記憶部103に記憶させるとともに、一覧表作成部1204に起動指示をする。

【0052】即ち、インデックス分割部1202から最初に一つのインデックス番号1303の例えば「106」の通知を受けると、対応するポインタ1302の値「18204」をインデックスファイル一時記憶部102から読み出し、その値を「106」で除算した商を四捨五入した整数値「172」を計算し、この整数値を傾きとする一次関数「 $Y=172X$ 」を予測関数として求め、予測関数記憶部103に図17に示すように記憶させる。

【0053】次に、インデックス分割部1202から一つまたはそれ以上のインデックス番号1303の通知を受けると、通知を受けた若いインデックス番号1303の順にそのインデックス番号1303を区間の最後とする区間に分割して、上記と同様に新たに予測関数を求め、予測関数記憶部103の記憶内容を更新する。即ち、インデックス分割部1202からインデックス番号1303の「47」の通知を受けると、インデックス番号1303の「0」～「47」までの区間と「48」～「106」までの区間の予測関数を求める。インデックス番号1303の「47」のポインタ1302の値「816」をインデックスファイル一時記憶部102から読み出し、その値「47」で除算した商を四捨五入した整数値「17」から一次関数「 $Y=17X$ 」をインデックス番号「0」～「47」の区間の予測関数として求める。次に、インデックス番号1303の「48」のポインタ1302の値「1116」を読み出し、インデックス番号1303の「106」の「18204」から $(18204-1116)/(106-48)$ の計算をし、その四捨五入した整数値「295」を一次関数の傾きとして得る。この傾きからポインタ1302の値「1116」とインデックス番号1303の「48」とからインデックス番号「48」～「106」の区間の予測関数「 $Y=295X-13044$ 」を求める。得られた2つの予測関数を予測関数記憶部103に図18に示すように記憶させる。

【0054】一覧表作成部1204は、関数作成部1203から起動指示を受けると、インデックス分割部1202が分割したインデックスファイル中の各区間と、予

測関数記憶部103に記憶された関数とを組として一覧表を作成し、一覧表一時記憶部1205に記憶させる。即ち、最初に関数作成部1203から起動指示を受けたときは、区間は1つで、これに対応する関数は $Y=172X$ である。この区間と関数の組を一覧表一時記憶部1205に記憶させるとともに関数計算部1206に起動指示をする。図19は、この状態での各一時記憶部102、107、109、1205の内容を表している。

【0055】また、一覧表作成部1204は、関数作成部1203から起動指示を再度受けると、インデックス分割部1202が分割したインデックスファイル中の各区間と、予測関数記憶部103に新たに記憶された関数とを組として一覧表を作成し、一覧表一時記憶部1208に記憶させる。即ち、区間は2つで、2つの区間に対応する関数は $Y=17X$ と $Y=295-13044$ である。この区間と関数との組を一覧表一時記憶部1208に記憶させるとともに関数計算部1206に起動指示をする。図20は、この状態での各一時記憶部102、107、109、1205の内容を表している。

【0056】一覧表一時記憶部1205は、高速半導体メモリ等で実現され、一覧表作成部1204が作成した一覧表を例えば図19、20等のように記憶している。図19に示す記憶内容では、全区間に対して関数 $Y=172X$ が適用され、図20に示す記憶内容では、インデックス番号0から47までの区間に対して $Y=17X$ 、インデックス番号48から106の区間に対して関数 $Y=295X-13044$ がそれぞれ適用される。

【0057】関数計算部1206は、一覧表作成部1204から起動指示を受けて、一覧表一時記憶部1205に記憶されている予測関数を読み出し、インデックスファイル一時記憶部102に記憶されているポインタ1302に対応するインデックス番号1303を入力変数 X に代入して、予測ポインタの値である出力変数 Y を計算し、予測ポインタ一時記憶部107に記憶させる。予測ポインタの計算が終了すると、差分計算部108に起動指示をする。

【0058】図19に示したように、インデックスファイル一時記憶部102に各ポインタ1302が記憶され、一覧表一時記憶部1205に予測関数 $Y=172X$ が記憶されているときは、インデックス番号1303の「0」から「106」まで順次予測ポインタの値を計算し、予測ポインタ一時記憶部1205に記憶させる。この結果は、図21に示すようになる。

【0059】図20に示したように、一覧表一時記憶部1205にインデックス番号1303の「0」から「47」の区間に予測関数 $Y=17X$ 及びインデックス番号1303の「48」から「106」の区間に予測関数 $Y=295X-13044$ が記憶されているときは、それぞれの区間にその予測関数を適用して、順次予測ポインタの値を計算し、予測ポインタ一時記憶部1205に

図22に示すように記憶させる。

【0060】差分計算部108は、上記第1実施例と同様に差分の計算を終了すると、分割指示部1207に起動指示をする。差分一時記憶部109は、図21に示す状態の後、差分計算部108の計算結果を図23に示すように記憶し、図22に示す状態の後、差分計算部108の計算結果を図24に示すように記憶している。

【0061】分割指示部1207は、差分計算部108から起動指示を受けると、差分範囲記憶部1201に記憶された差分の下限と上限との値を読み出し、差分一時記憶部109に記憶されている差分が差分の下限と上限との範囲内にあるか否かを全差分について判定する。この差分の下限と上限との範囲内にあるときは一覧表出力部1208に起動指示をし、範囲外にあるときは予測ポインタ一時記憶部107と差分一時記憶部109と一覧表一時記憶部1205とに記憶されている内容を抹消する、とともにインデックス分割部1202に起動指示をする。

【0062】分割指示部1207は、差分一時記憶部109に図23に示す差分が記憶されているとき、差分の下限と上限との範囲「 $-128 \sim 127$ 」外となると判定する。この場合には、区間を分割して予測関数を新たに作成するためインデックス分割部1202に指示が与えられる。差分一時記憶部109に図24に示す差分が記憶されているとき、全ての差分は、差分の下限と上限の範囲内であると判定する。この場合には、差分一時記憶部109に記憶されている全ての差分は、各々8ビットの記憶容量の範囲で記録することができる。

【0063】一覧表出力部1208は、分割指示部1207から起動指示を受けると、差分一時記憶部109に記憶された差分と一覧表一時記憶部1205に記憶された一覧表とをそれぞれ読み出し、差分インデックスファイルを記録する記録部（図示せず）に出力する。差分は、図25に示すように、差分一時記憶部109に記憶されている差分をバイナリ表現して出力される。一覧表は、図26に示すように、区間の最初と最後のインデックス番号と対応する予測関数とを組にした集合として出力される。なお、予測関数は、そのまま出力されてもよいけれども、予測関数が一次関数の場合には、その係数である傾きと切片とだけを出力することも可能である。また、差分と一覧表とは、差分インデックスファイルを記録する記録部に出力されたけれども、差分は、辞書データとともに差分インデックスファイルとして例えばCD-ROMに記録するようにし、一覧表は、そのCD-ROMの再生装置のプログラムとして記録されるようにしてもよい。

【0064】次に、本実施例の動作について図27のフローチャートを用いて説明する。入力部101は、ポインタ1302が格納されているインデックスファイルが入力されると（S2702）、インデックス番号130

3の順にポインタ1302を順次インデックスファイル一時記憶部102に記憶させる(S2704)。インデックス分割部1202は、インデックスファイルの最後のインデックス番号を取得し(S2706)、関数作成部1203に通知する(S2708)。

【0065】関数作成部1203は、通知されたインデックス番号を境界として分割区間を決定し、ポインタを予測するための関数を作成する(S2710)。作成した予測関数を予測関数記憶部103に記憶させる(S2712)。一覧表作成部1204は、分割区間と予測関数の一覧表を作成する(S2714)。作成した一覧表を一覧表一時記憶部1205に記憶させる(S2716)。

【0066】関数計算部1206は、一覧表一時記憶部1205から予測関数を取得する(S2718)。取得した関数にインデックス番号を代入して予測ポインタを計算する(S2720)。予測ポインタを予測ポインタ一時記憶部107に記憶させる(S2722)。差分計算部108は、インデックスファイル一時記憶部102に記憶されているポインタの値と予測ポインタ一時記憶部107に記憶されている予測ポインタの値との差分を計算する(S2724)。計算した差分を差分一時記憶部109に記憶させる(S2726)。

【0067】分割指示部1207は、全ての差分が差分範囲記憶部1201に記憶されている差分の下限と上限との範囲内にあるか否かを判定する(S2728)。範囲内ないと判定されたときは、インデックス分割部1202は、インデックスファイル一時記憶部102に記憶されているポインタの増加値の変化値を計算し(S2730)、差分範囲記憶部1201に記憶された差分範囲外の変化値に対応するインデックス番号を取得し(S2732)、S2708に戻る。

【0068】差分が範囲内にあると判定されたときは、一覧表出力部1208は、差分一時記憶部109に記憶されている差分を読み出して出力し(S2734)、一覧表一時記憶部1205に記憶されている一覧表を読み出して出力して(S2736)、処理を終了する。

(第3実施例) 図28は、本発明に係るインデックス伸長装置の第3実施例の構成図である。

【0069】インデックス伸長装置は、差分インデックスファイルを入力する入力部2801と、高速半導体メモリ等からなる差分インデックスファイル一時記憶部2802と、データ要素検索のためのインデックス番号指定部2803と、予測関数記憶部2804と、予測ポインタの値を計算する関数計算部2805と、高速半導体メモリ等からなる予測ポインタ一時記憶部2806と、ポインタ計算部2807と、高速半導体メモリ等からなるポインタ一時記憶部2808と、出力部2809とを備える。

【0070】入力部2801は、図29に示す電子英和

辞書の差分インデックスファイル2901を読み込み、差分インデックスファイル一時記憶部2802に一時記憶させるとともに関数計算部2805に起動指示をする。ここで、図29に示す電子英和辞書は、バイナリ表現された差分インデックスファイル2901と、英和辞書データ2902とを有している。差分インデックスファイル2901は、10進表現された差分インデックスファイル2903と等価である。差分インデックスファイル2903の左方の数字は、各差分インデックス2904に対応するインデックス番号2905を示している。この差分インデックス2904は、所定の手順で英和辞書データ2902のデータ要素のポインタ2906に復元される。各ポインタ2906は、矢印2907で示すように各データ要素の辞書データファイルの格納位置を示している。

【0071】図30は、差分インデックスファイル2901の入力直後の各一時記憶部2802、2806、2808の内容を表したものであり、差分インデックスファイル一時記憶部2802に各差分インデックス2904が記憶されている。インデックス番号指定部2803は、ユーザの英和辞書の検索の際に、差分インデックス2904を辞書データ要素のポインタ2906に復元するためにインデックス番号2905を指定されると、関数計算部2805にそのインデックス番号2905を通知する。

【0072】予測関数記憶部2804は、図31に示すように、差分インデックス2904に対応するポインタ2906の値を予測する予測関数を予め記憶している。予測関数 $Y=21X$ の入力変数 X にインデックス番号2905を入力すると出力変数 Y が予測ポインタとして計算される。なお、この予測関数は、上記第1又は第2実施例のインデックス圧縮装置で用いられた予測関数に対応するものである。

【0073】関数計算部2805は、インデックス番号指定部2803からインデックス番号2905の通知を受けると、予測関数記憶部2804に記憶されている予測関数を読み出し、その予測関数に通知されたインデックス番号2905を代入して予測ポインタを計算する。計算した予測ポインタを予測ポインタ一時記憶部2806に一時記憶させるとともにポインタ計算部2807にインデックス番号を通知する。例えば、インデックス番号指示部2803からインデックス番号2905の

「2」の通知を受けると、予測関数記憶部2804から予測関数 $Y=21X$ を読み出し、 $Y=21 \times 2=42$ を計算し、予測ポインタ一時記憶部に「42」を記憶させる。

【0074】予測ポインタ一時記憶部2806は、図32に示すように、関数計算部2805で計算された予測ポインタをインデックス番号2905に対応して記憶している。ポインタ計算部2807は、関数計算部280

5からインデックス番号2905の通知を受けると、差分インデックスファイル一時記憶部2802に記憶されているそのインデックス番号2905に対応する差分インデックスと予測ポインタ一時記憶部2806に記憶されている予測ポインタとの和を計算し、ポインタ一時記憶部2808に一時記憶させるとともに出力部2809に起動指示をする。例えば、図32に示すように、インデックス番号2905の「2」の差分インデックス2904の値「-1」と予測ポインタの値「42」の和は、「41」と計算される。

【0075】ポインタ一時記憶部2808は、ポインタ計算部2807で計算されたポインタの値を図33に示すように記憶している。出力部2809は、ポインタ計算部2807から起動指示を受けると、ポインタ一時記憶部2808に記憶されているポインタを読み出して辞書検索部（図示せず）に出力する。

【0076】次に本実施例の動作を図34のフローチャートを用いて説明する。入力部2801は、差分インデックス2904を格納している差分インデックスファイル2903の入力を受け付けると（S3402）、差分インデックスファイル一時記憶部2802に記憶させる（S3404）。インデックス番号指定部2803でインデックス番号が指定されると（S3406）、関数計算部2805は、予測関数記憶部2804から予測関数を取得し（S3408）、指定されたインデックス番号を予測関数に代入して予測ポインタを計算する（S3410）。計算した予測ポインタを予測ポインタ一時記憶部2806に記憶させる（S3412）。

【0077】ポインタ計算部2807は、差分インデックスファイル一時記憶部2802に記憶されている差分インデックス2904の値と予測ポインタ一時記憶部2806に記憶されている予測ポインタの値との和であるポインタの値を計算し（S3414）、計算したポインタをポインタ一時記憶部2808に記憶させる。出力部2809は、ポインタ一時記憶部2808に記憶されているポインタを読み出し、出力する（S3418）。

【0078】なお、本実施例において予測関数記憶部2804は、予測関数を予め記憶しているとしたけれども、他の実施例として、図10に示す差分インデックスファイル1001にヘッダ情報1004を有している場合には、ヘッダ情報に基づいて予測関数を作成する予測関数作成部を設けて、この予測関数記憶部2804に換えてもよい。

【0079】また、本実施例では、予測関数記憶部2804は一の予測関数を記憶しているとしたけれども、第2実施例の図26に示したように、複数の関数を一覧表として記憶している場合でも、同様にポインタを復元できることはいうまでもない。

【0080】

【発明の効果】以上説明したように、請求項1の発明に

よれば、省記憶容量のインデックスファイルとすることができるので、低コストの電子辞書を得ることができる。請求項2の発明によれば、複数の系列データを含む電子辞書のインデックスファイルを省記憶容量とすることができる。

【0081】請求項3の発明によれば、インデックスファイルのポインタを差分を用いて省記憶容量とした電子辞書を得ることができる。請求項4の発明によれば、複数の系列データを含む場合にも省記憶容量のインデックスファイルを有する電子辞書を得ることができる。請求項5の発明によれば、入力されたインデックスファイル中のポインタに対して、関数を用いてこのポインタの値を予測する予測ポインタを計算して、入力されたポインタと予測ポインタとの差を計算し、この差をポインタの代わりにインデックスとすることによってインデックスを圧縮することができる。

【0082】請求項6の発明によれば、入力されたインデックスファイル中のポインタの値に応じて、予測ポインタの計算に用いる関数を使い分け、計算された差分と計算に用いた関数を表す識別子を組として出力することにより、複数の系列データが混在するインデックスファイルを圧縮することができる。請求項7の発明によれば、ポインタの値を予測するために用いる関数を1次関数に限定することにより、傾きと切片だけを記憶するだけでよく、定数記憶手段の記憶領域を削減することができる。

【0083】請求項8の発明によれば、ポインタの値を予測する関数をインデックスファイルのポインタの値から決定することができ、かつ、ポインタに代えて用いる差分を記憶する領域のビット数を差分範囲の設定によって選ぶことができる。請求項9の発明によれば、圧縮されたインデックス中の差分に、予測ポインタの値を加えポインタの値を復元し、圧縮されたインデックスでの元のインデックスを用いたデータの検索を実施可能としている。

【0084】請求項10の発明によれば、複数の関数で差分インデックスが計算されていた場合にも、ポインタを容易に復元することができる。

【図面の簡単な説明】

【図1】本発明に係るインデックス圧縮装置の第1実施例の構成図である。

【図2】本実施例における圧縮をするインデックスファイルとそれを含む英和辞書の一例を示す図である。

【図3】図2に示すインデックスファイルバイナリ表示を示す図である。

【図4】本実施例におけるインデックスファイルの入力直後の各一時記憶部の記憶内容を示した図である。

【図5】本実施例における予測関数記憶部の記憶内容を示した図である。

【図6】本実施例における各一時記憶部の記憶内容を示

した図である。

【図 7】本実施例における各一時記憶部の記憶内容を示した図である。

【図 8】本実施例における各一時記憶部の記憶内容を示した図である。

【図 9】本実施例における差分一時記憶部と関数識別子一時記憶部の記憶内容と出力される差分インデックスファイルの内容を示した図である。

【図 10】本発明に係る電子辞書の一例を示す図である。

【図 11】本実施例の動作を説明するフローチャートである。

【図 12】本発明に係るインデックス圧縮装置の第 2 実施例の構成図である。

【図 13】本実施例における圧縮をするインデックスファイルとそれを含む固有名詞辞書の一例を示す図である。

【図 14】図 13 に示すインデックスファイルのバイナリ表示を示す図である。

【図 15】本実施例における各一時記憶部の記憶内容を示した図である。

【図 16】本実施例における差分範囲記憶部の記憶内容を示した図である。

【図 17】本実施例における予測関数記憶部の記憶内容を示した図である。

【図 18】本実施例における予測関数記憶部の記憶内容を示した図である。

【図 19】本実施例における各一時記憶部の記憶内容を示した図である。

【図 20】本実施例における各一時記憶部の記憶内容を示した図である。

【図 21】本実施例における各一時記憶部の記憶内容を示した図である。

【図 22】本実施例における各一時記憶部の記憶内容を示した図である。

【図 23】本実施例における各一時記憶部の記憶内容を示した図である。

【図 24】本実施例における各一時記憶部の記憶内容を示した図である。

【図 25】本実施例における差分一時記憶部の記憶内容と出力される差分インデックスファイルの内容を示した図である。

【図 26】本実施例における出力される一覧表を示した図である。

【図 27】本実施例の動作を説明するフローチャートである。

【図 28】本発明に係るインデックス伸長装置の第 3 実施例の構成図である。

【図 29】本実施例における伸長をする差分インデックスファイルとそれを含む英和辞書の一例を示す図である。

【図 30】本実施例における各一時記憶部の記憶内容を示した図である。

【図 31】本実施例における予測関数記憶部の記憶内容を示した図である。

10 【図 32】本実施例における各一時記憶部の記憶内容を示した図である。

【図 33】本実施例における各一時記憶部の記憶内容を示した図である。

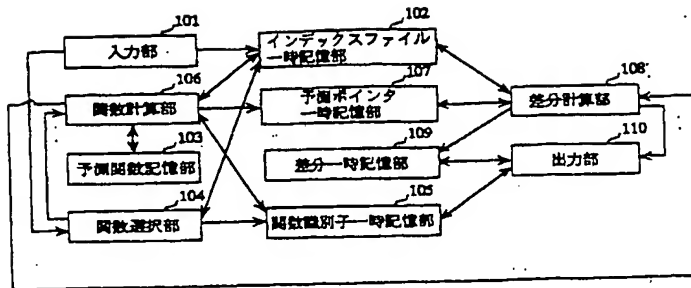
【図 34】本実施例の動作を説明するフローチャートである。

【図 35】従来のインデックスファイルを含む英和辞書を示す図である。

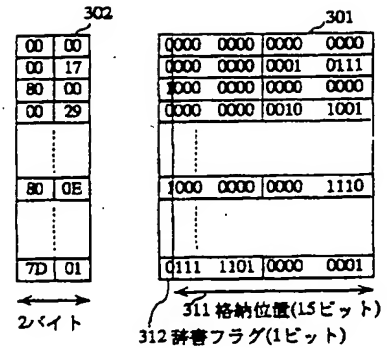
【符号の説明】

- 101 入力部
- 102 インデックスファイル一時記憶部
- 103 予測関数記憶部
- 104 関数選択部
- 105 関数識別子一時記憶部
- 106 関数計算部
- 107 予測ポインター一時記憶部
- 108 差分計算部
- 109 差分一時記憶部
- 110 出力部
- 1001 差分インデックスファイル
- 1201 差分範囲記憶部
- 1202 インデックス分割部
- 1203 関数作成部
- 1204 一覧表作成部
- 1205 一覧表一時記憶部
- 1206 関数計算部
- 1207 分割指示部
- 1208 一覧表出力部
- 2801 入力部
- 2802 差分インデックスファイル一時記憶部
- 2803 インデックス番号指定部
- 2804 予測関数記憶部
- 2805 関数計算部
- 2806 予測ポインター一時記憶部
- 2807 ポインタ計算部
- 2808 ポインター一時記憶部
- 2809 出力部

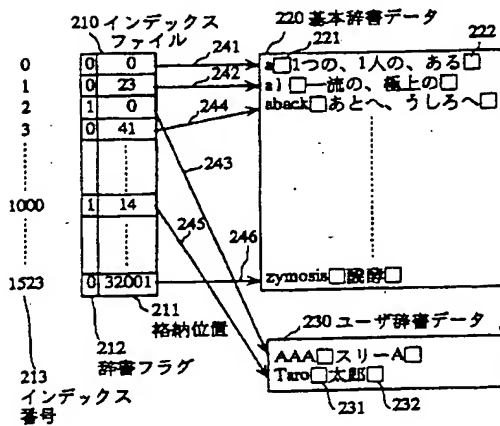
【図1】



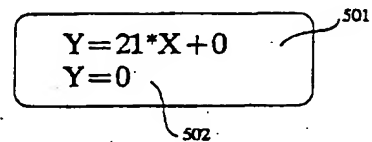
【図3】



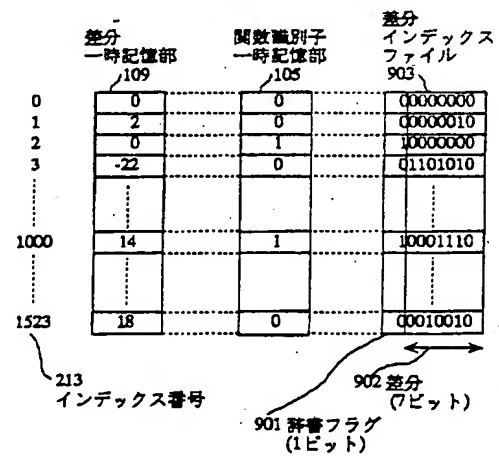
【図2】



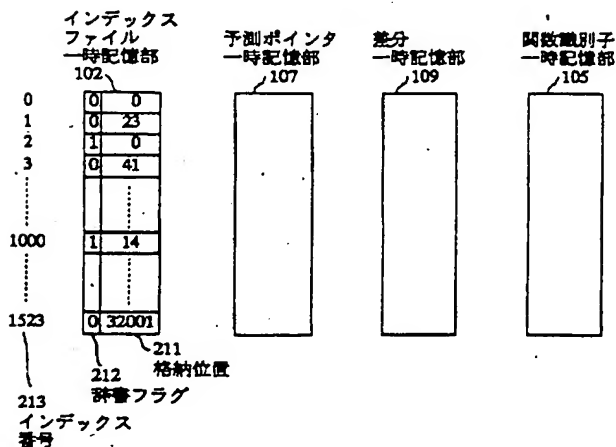
【図5】



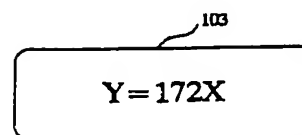
【図9】



【図4】



【図17】



【図6】

インデックス ファイル 一時記憶部 102		予測ポインタ 一時記憶部 107	差分 一時記憶部 109	関数識別子 一時記憶部 105
0	0			0
1	23			0
2	1			1
3	0			0
...
1000	1			1
...
1523	0			0

211 格納位置
212 辞書フラグ
213 インデックス
番号

【図13】

1301 インデックス ファイル		1304 固有名詞辞書データ
0	0	北海道□都道府県名、地名□
1	24	青森□県名、地名□
2	40	岩手□県名、地名□
...
46	800	沖縄□県名、地名□
47	816	西郷隆盛□薩摩出身□
48	1116	伊藤博文□初代総理□
...
106	18204	徳川家康□江戸幕府将軍□

1302 ポインタ
1303 インデックス番号

【図14】

【図7】

インデックス ファイル 一時記憶部 102		予測ポインタ 一時記憶部 107	差分 一時記憶部 109	関数識別子 一時記憶部 105
0	0	0		0
1	0	23		0
2	1	0		1
3	0	41		0
...
1000	1	14		1
...
1523	0	32001		0

211 格納位置
212 辞書フラグ
213 インデックス
番号

1401 バイナリ インデックス ファイル	
0	00 00
1	00 18
2	00 28
...	...
46	03 20
47	03 30
48	04 5C
...	...
106	47 1C

下位1バイト
上位1バイト
1303 インデックス番号

【図25】

【図16】

下限 : -128
上限 : +127

【図18】

$Y = 17X$
 $Y = 295X - 13044$

差分 一時記憶部 109	出力
0	00000000
1	00000111
2	00000110
...	...
46	00001000
47	00010001
48	00000000
...	...
106	00010110

1303 インデックス番号

【図8】

インデックス ファイル 一時記憶部	予測ポイント 一時記憶部	差分 一時記憶部	関数識別子 一時記憶部
102	107	109	105
0 0	0	0	0
1 23	21	2	0
2 0	0	0	1
3 41	63	-22	0
...
1000 1 14	0	14	1
...
1523 0 32001	31983	18	0

211 格納位置
212 辞書フラグ
213 インデックス
番号

【図30】

差分インデックス ファイル 一時記憶部	予測ポイント 一時記憶部	ポイント 一時記憶部
2802	2806	2808
0		
1 2		
2 -1		
...		
1523 18		

2904 差分インデックス
2905 インデックス番号

【図10】

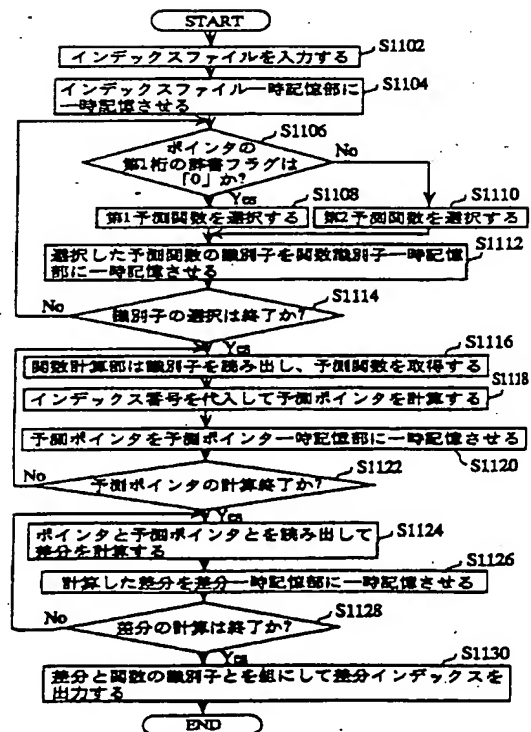
【図11】

1001 差分インデックスファイル	1002 基本辞書データ	1003 ユーザ辞書データ
関数値数 2	関数0の値 21	関数0の値 21
関数0の切片 0	関数1の値 0	関数1の値 0
関数1の切片 0	差分値数 1524	差分値数 1524
差分値数 1524	フラグビット数 7	フラグビット数 7
...
1000 1 14
1523 0 18

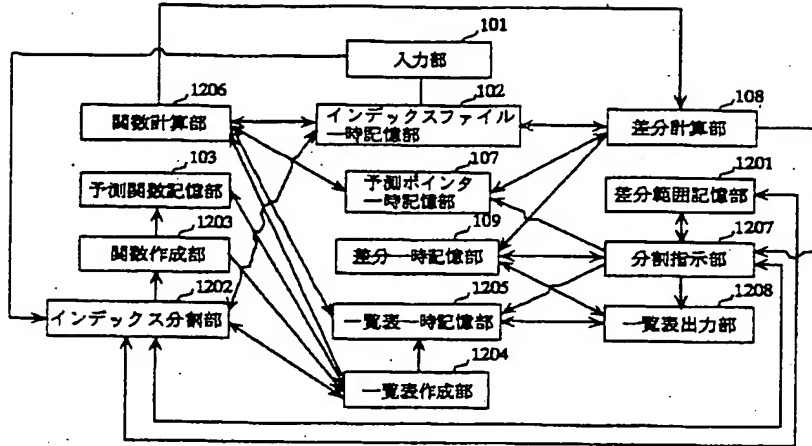
1004 ヘッダ情報
1005 辞書フラグ
1006 差分
1007 関数値数
1008 関数0の値
1009 関数0の切片
1010 関数1の値
1011 関数1の切片
1012 差分値数
1013 インデックス
番号

【図26】

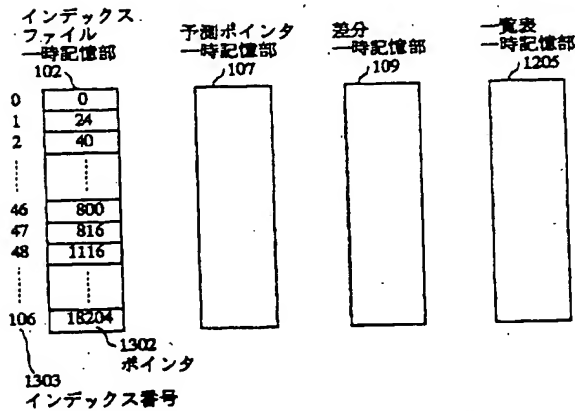
インデックス番号	関数
0~47	$Y=17X$
48~106	$Y=295X-13044$



【図12】



【図15】

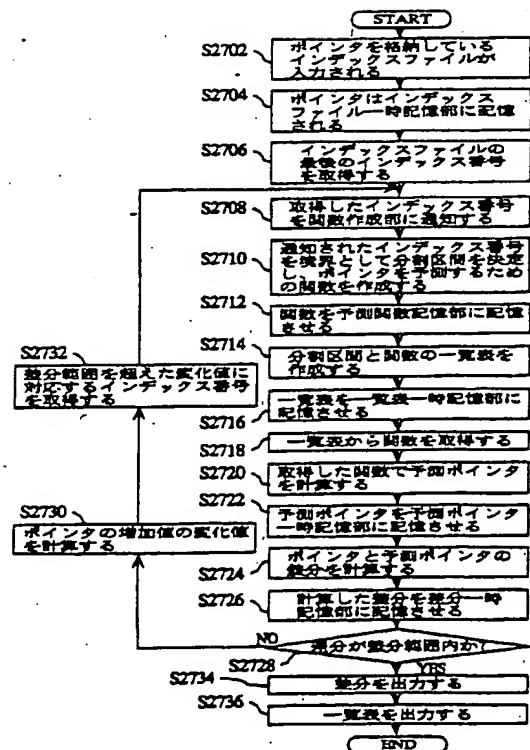


【図31】

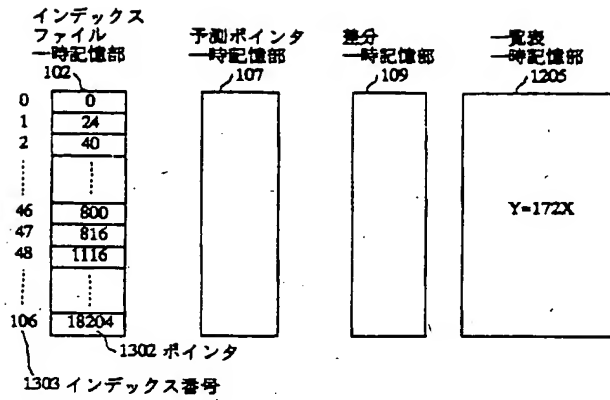
2804

$$Y = 21 \times X$$

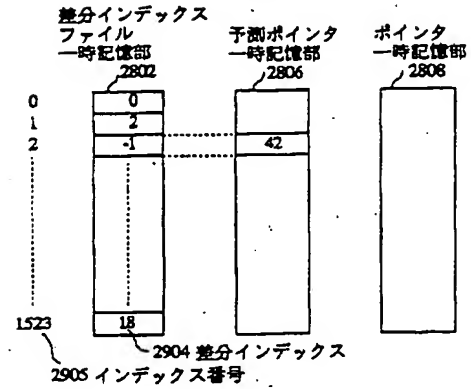
【図27】



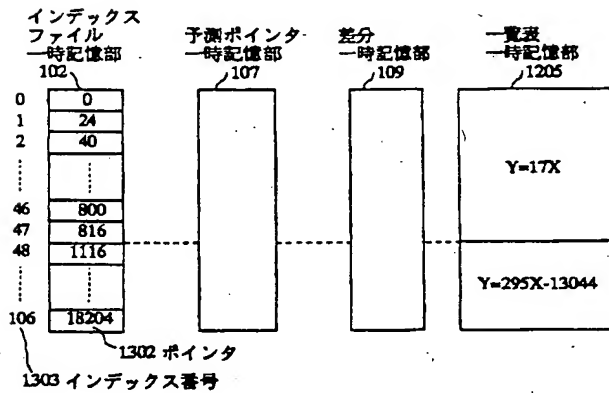
【図19】



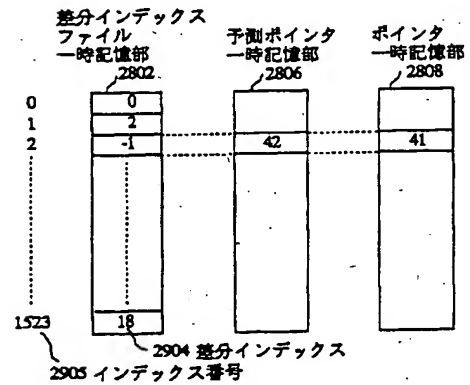
【図32】



【図20】

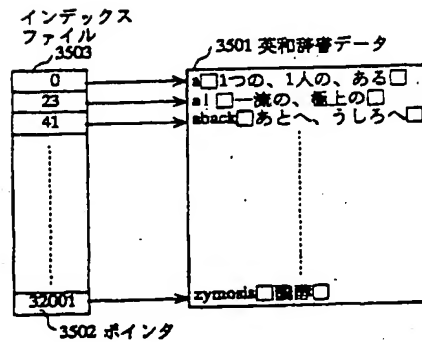
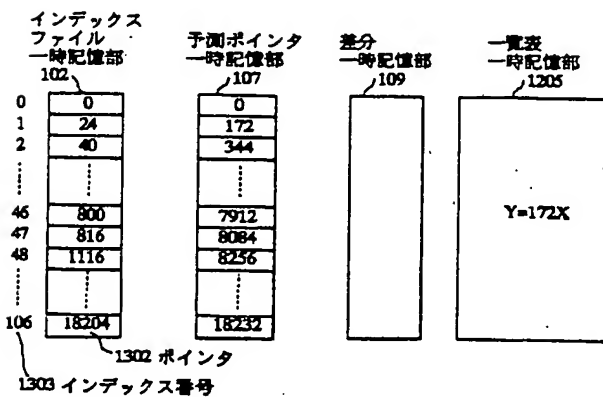


【図33】

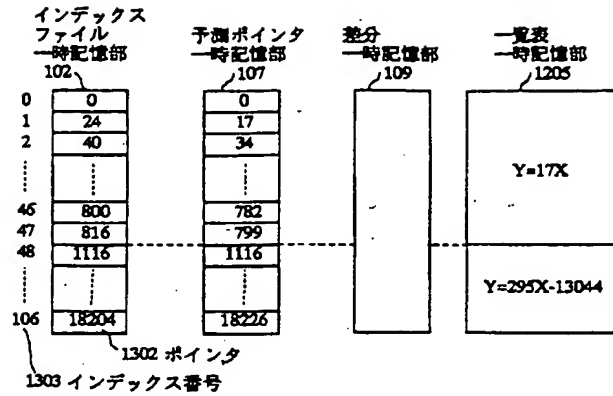


【図35】

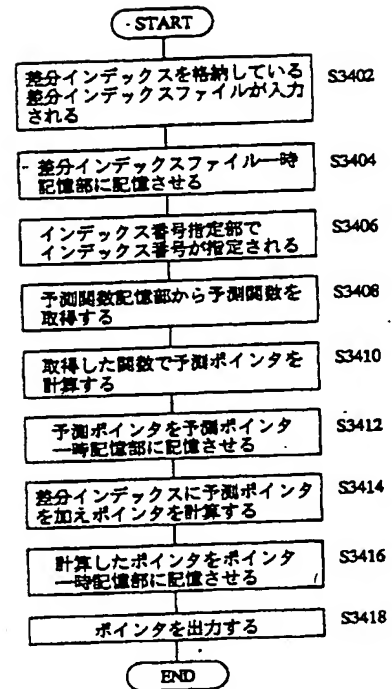
【図21】



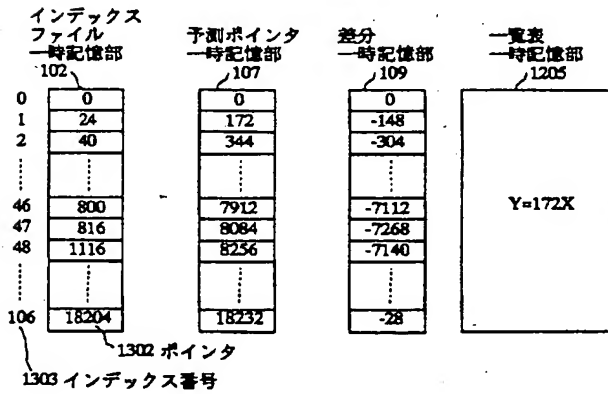
【図22】



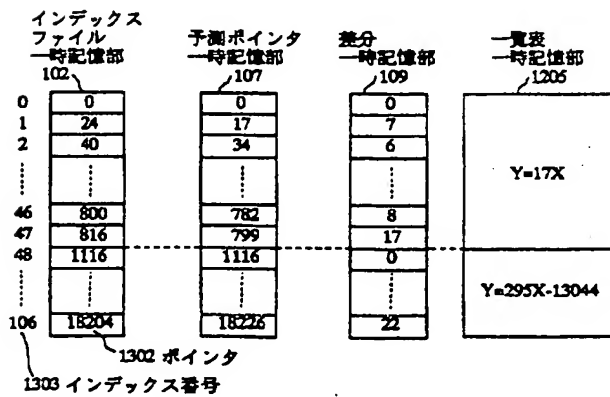
【図34】



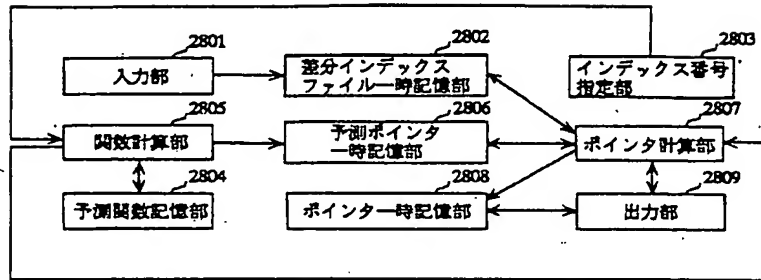
【図23】



【図24】



【図28】



【図29】

